

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-222342

(43) 公開日 平成10年(1998) 8月21日

(51) Int.Cl.<sup>6</sup>

識別記号

F I

G 0 6 F 3/16

3 2 0

G 0 6 F 3/16

3 2 0 B

12/00

5 4 7

12/00

5 4 7 H

審査請求 未請求 請求項の数 5 O L (全 6 頁)

(21) 出願番号 特願平9-24024

(22) 出願日 平成9年(1997) 2月6日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 淵 武志

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72) 発明者 加藤 恒昭

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(74) 代理人 弁理士 伊東 忠彦

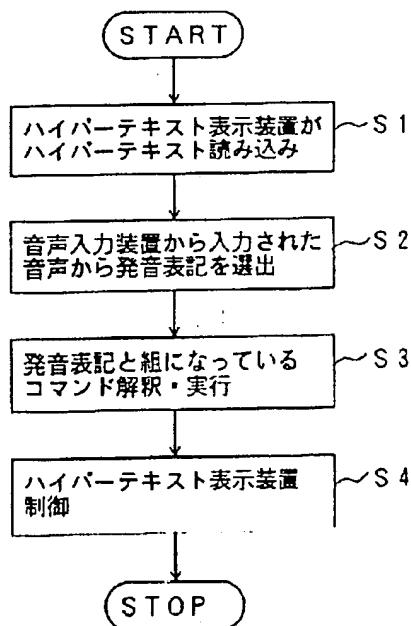
(54) 【発明の名称】 ハイパーテキスト音声制御方法及び装置

(57) 【要約】

【課題】 音声認識の対象語及びそれに結びつく処理をハイパーテキスト中で指定することが可能とし、柔軟にハイパーテキスト表示装置を音声制御することが可能なハイパーテキスト音声制御方法及び装置を提供する。

【解決手段】 本発明は、ハイパーテキスト表示装置がハイパーテキストを要求し、発音表記とコマンドの組からなるリストを含むハイパーテキストをハイパーテキスト表示装置が読み込むと、音声入力装置によって入力された音声に対応する発音表記を選出し、発音表記と組になっているコマンドを解釈し、実行することによりハイパーテキスト表示装置を制御する。

本発明の原理を説明するための図



BEST AVAILABLE COPY

## 【 特許請求の範囲】

【 請求項1 】 ハイパーテキスト 表示装置を音声によって制御するためのハイパーテキスト 音声制御方法において、

ハイパーテキスト 中に記述された発音表記とコマンドの組を用いて動的に音声認識制御機能を変化させて前記ハイパーテキスト 表示装置を制御することを特徴とするハイパーテキスト 音声制御方法。

【 請求項2 】 発音表記とコマンドの組からなるリストを含むハイパーテキストを前記ハイパーテキスト 表示装置が読み込むと、

音声入力装置によって入力された音声に対応する発音表記を選出し、

前記発音表記と組になっているコマンドを解釈し、実行することにより前記ハイパーテキスト 表示装置を制御する請求項1 記載のハイパーテキスト 音声制御方法。

【 請求項3 】 前記ハイパーテキスト 表示装置は、コンピュータネットワークまたはファイルシステムから、音声認識機能を示すタグ及び発音表記とコマンドの組からなる前記リストを含むハイパーテキストを読み込む請求項2 記載のハイパーテキスト 音声制御方法。

【 請求項4 】 通常のハイパーテキスト 表示機能に加え、ハイパーテキストを読み込む際に、該ハイパーテキスト 中の特定のタグと該タグに続く発音表記とコマンドの組からなるリストを検出するハイパーテキスト 表示装置と、

利用者が発声した音声を入力する音声入力手段と、発音表記とコマンドの組からなるリストを前記ハイパーテキスト 表示装置から取得し、該発音表記を以降の認識語彙とし、前記音声入力手段から得られた音声該認識語彙中のどれに最も近いかを判定し、認識結果として該発音表記と組になっているコマンドを返す音声認識手段と、

前記音声認識手段によって得られた前記コマンドを解釈し、解釈結果に基づいて前記ハイパーテキスト 表示装置を制御するコマンド 解釈実行手段とを有することを特徴とするハイパーテキスト 音声制御装置。

【 請求項5 】 前記ハイパーテキスト 表示装置は、コンピュータネットワークまたはファイルシステムから、音声認識機能を示すタグ及び発音表記とコマンドの組からなる前記リストを含むハイパーテキストを読み込む請求項4 記載のハイパーテキスト 音声制御装置。

## 【 発明の詳細な説明】

## 【 0 0 0 1 】

【 発明の属する技術分野】本発明は、ハイパーテキスト 音声制御方法及び装置に係り、特に、ハイパーテキスト 表示装置を音声によって制御するための、ハイパーテキスト 中の記述に応じて、自動的に制御動作を変えるハイパーテキスト 音声制御方法及び装置に関する。

## 【 0 0 0 2 】

【 従来の技術】インターネットなどのコンピュータネットワークを用いて、HTML やSGML などのハイパーテキストを受信して、表示することが有力な情報収集手段となっている。その際に、一般にブラウザと呼ばれるハイパーテキスト 表示装置が用いられる。このブラウザを音声によって制御したいというニーズがある。

【 0 0 0 3 】従来の技術では、音声認識させたい発音表記と対応するコマンドとが固定されている。そして、その発音表記の中の一つが認識されると、それに対応するコマンドが実行されるという形で、ブラウザの音声による制御が実現されている。

## 【 0 0 0 4 】

【 発明が解決しようとする課題】しかしながら、ハイパーテキストを用いて情報を提供しようとする場合、様々な処理を音声に反応させたいというニーズがある。しかし、従来の技術では、音声認識の対象語と、それに結びついた処理が固定されていたため、様々なニーズにあった形で音声認識機能を利用することができないという問題がある。

【 0 0 0 5 】本発明は、上記の点に鑑みなされたもので、音声認識の対象語及びそれに結びつく処理をハイパーテキスト 中で指定することが可能とし、柔軟にハイパーテキスト 表示装置を音声制御することが可能なハイパーテキスト 音声制御方法及び装置を提供することを目的とする。

## 【 0 0 0 6 】

【 課題を解決するための手段】図1 は、本発明の原理を説明するための図である。本発明は、ハイパーテキスト 表示装置を音声によって制御するためのハイパーテキスト 音声制御方法において、ハイパーテキスト 中に記述された発音表記とコマンドの組を用いて動的に音声認識制御機能を変化させてハイパーテキスト 表示装置を制御する。

【 0 0 0 7 】また、本発明は、ハイパーテキスト 表示装置がハイパーテキストを要求し、発音表記とコマンドの組からなるリストを含むハイパーテキストをハイパーテキスト 表示装置が読み込むと（ステップ1）、音声入力装置によって入力された音声に対応する発音表記を選出し（ステップ2）、発音表記と組になっているコマンドを解釈し、実行する（ステップ3）ことによりハイパーテキスト 表示装置を制御する（ステップ4）。

【 0 0 0 8 】また、本発明は、ハイパーテキスト 表示装置は、コンピュータネットワークまたはファイルシステムから、音声認識機能を示すタグ及び発音表記とコマンドの組からなるリストを含むハイパーテキストを読み込む。図2 は、本発明の原理構成図である。本発明のハイパーテキスト 音声制御装置は、通常のハイパーテキスト 表示機能に加え、ハイパーテキストを読み込む際に、該ハイパーテキスト 中の特定のタグと該タグに続く発音表記とコマンドの組からなるリストを検出するハイパーテ

3

キスト表示装置20と、利用者が発声した音声を入力する音声入力手段30と、発音表記とコマンドの組からなるリストをハイパーテキスト表示装置20から取得し、該発音表記を以降の認識語彙とし、音声入力手段から得られた音声の該認識語彙中のどれに最も近いかを判定し、認識結果として該発音表記と組になっているコマンドを返す音声認識手段40と、音声認識手段40によって得られたコマンドを解釈し、解釈結果に基づいてハイパーテキスト表示装置20を制御するコマンド解釈実行手段50とを有する。

【0009】また、上記のハイパーテキスト表示装置20は、コンピュータネットワークまたはファイルシステムから、音声認識機能を示すタグ及び発音表記とコマンドの組からなるリストを含むハイパーテキストを読み込む。本発明では、コンピュータネットワークまたは、ファイルシステムから音声認識機能を示すタグ（以下、音声認識タグと記す）及び発音表記とコマンドの組のリストが書き込まれたハイパーテキストを、ハイパーテキスト表示装置が読み込む。このハイパーテキスト表示装置20は、音声認識タグを検出すると、発音表記とコマンドの組のリストを音声認識手段40に渡す。音声認識手段40は、音声入力手段30に音声データを要求する。音声入力手段30から音声データが音声認識手段に渡されると、音声認識手段40は音声認識処理を行い、先に受け取った発音表記の中からその音声データに最も近い発音表記を選択する。そして、選択した発音表記と組になっているコマンドをコマンド解釈実行手段50に渡す。コマンド解釈実行手段50は、そのコマンドを解釈した結果に基づいて様々なハイパーテキスト表示装置20の動作制御を行う。

【0010】これにより、ハイパーテキスト中に音声表記とコマンドを指定することによって、任意の処理を音声と結び付けて実行させることが可能となる。

【0011】

【発明の実施の形態】図3は、本発明のハイパーテキスト音声制御装置の構成を示す。同図に示すハイパーテキスト音声制御装置は、ハイパーテキスト表示装置20、音声入力装置30、音声認識装置40、コマンド解釈実行装置50から構成され、ハイパーテキスト表示装置20は、コンピュータネットワーク10に接続される。

【0012】ハイパーテキスト表示装置20は、音声認識タグ、発音表記とコマンドの組のリストが記述されたハイパーテキストを、コンピュータネットワーク10（または、ファイルシステム）から読み込む。当該ハイパーテキストには、音声認識タグが記述されている。ハイパーテキスト表示装置20は、この音声認識タグに引き続いて記述されている発音表記とコマンドの組のリストを音声認識装置40に渡す。

【0013】音声入力装置30はマイク等の機器から音声を入力し、音声認識装置40に音声データを渡す。音

4

声認識装置40は、発音表記とコマンドの組のリストを記録して、音声入力装置30からの入力を待つ。音声入力装置30から受け取った音声データを音声認識し、ハイパーテキスト表示装置20から受け取った発音表記の中で、その音声データと最も近い発音表記を選出する。音声認識装置40は、選出した発音表記と組になっていたコマンドをコマンド解釈実行装置50に送る。

【0014】コマンド解釈実行装置50は、音声認識装置40から受け取ったコマンドを解釈し、その解釈結果に基づいてハイパーテキスト表示装置20の表示切替等の様々な動作を実行する。図4は、本発明のハイパーテキスト音声制御動作を示す図である。

ステップ101）ハイパーテキスト表示装置20がネットワークまたは、ファイルシステムに、音声認識タグ及び発音表記とコマンドの組のリストが書き込まれたハイパーテキストの要求を発行する。

【0015】ステップ102）ハイパーテキスト表示装置20は、ハイパーテキスト、音声認識タグ及び発音表記とコマンドの組のリストを読み込む。

20 ステップ103）ハイパーテキスト表示装置20は、読み込んだハイパーテキストを表示する。

ステップ104）ハイパーテキスト表示装置20は、音声認識タグを検出すると、読み込んだ音声認識タグ及び発音表記とコマンドの組のリストを音声認識装置40に送る。

【0016】ステップ105）音声認識装置40は、音声入力装置30に音声データを要求する。

ステップ106）音声入力装置30は、音声データを音声認識装置40に渡す。

30 ステップ107）音声認識装置40は、音声データに対する音声認識処理を行い、先に受け取った発音表記の中からその音声データに最も近い発音表記を選択する。

【0017】ステップ108）音声認識装置40は、そして、選択した発音表記と組になっているコマンドをコマンド解釈実行装置50に渡す。

ステップ109）コマンド解釈実行装置50は、そのコマンドを解釈した結果に基づいてハイパーテキスト表示装置20に対して様々な動作制御を行う。

【0018】

40 【実施例】以下、図面と共に本発明の実施例を説明する。以下の実施例では、既存のハイパーテキスト表示装置に適用した場合の例を示す。この場合、ハイパーテキストはHTMLであり、ハイパーテキスト表示装置はブラウザと呼ばれる。本実施例では、"Netscape Communications Corporation"の提供する"Netscape Navigator"を例に説明する。

【0019】"Netscape Navigator"では、"Applet"と呼ばれる仕組みを用いることで、外部装置とのデータの受渡しが可能である。また、ここで用いたコマンドは、"JavaScript"と呼ばれるもので、コマンド解釈実行装置

50は、"Netscape Navigator"に内蔵されている。以下、図面と共に本発明の実施例を説明する。

【0020】図5は、本発明の一実施例のシステム構成図を示す。同図に示すシステムは、ブラウザ100、音声認識装置110、音声入力装置120及びマイク130から構成される。同図の構成において、ブラウザ100は、HTML表示装置101と"Java Script"解釈実行装置102を内蔵している。HTML表示装置101は図3におけるハイパーテキスト表示装置20に対応し、"Java Script"解釈実行装置102はコマンド解釈実行装置50に対応する。また、音声認識装置110は図3に示す音声認識装置40に、音声入力装置120は音声入力装置30に対応する。

【0021】図6は、本発明の一実施例の音声認識を可能とするHTMLの記述例である。ブラウザ100が、図6に示すようなHTMLを読み込むと、まず、HTML表示装置101が通常のHTMLの表示規則に従って文書を表示する。このとき、

<APPLET...

で示されるタグが検出されると、音声認識装置110とデータの受渡しが開始される。この例の場合、このタグが本発明における音声認識タグの役割を果たしている。発音表記とコマンドは、

<PARAM

で始まる部分に記述されている。図6の例では、『くだもの』

が発音表記であり、

window.open('http://www.cam/fruit.html')

がコマンドである。これは、括弧内で示されるHTMLにブラウザの表示を切り替える"Java Script"である。この発音表記とコマンドの組が音声認識装置110に渡される。音声入力装置120は、マイク130を介して音声を入力し、音声データとして音声認識装置110に渡す。

【0022】音声認識装置110は、このデータに対して音声認識を行い、HTML表示装置101から渡された発音表記の中で最も音声データに近いものを選択する。音声認識装置110は、選択した発音表記の組となるコマンドをJava Script 解釈実行装置102に渡す。Java Script 解釈実行装置102は、コマンドを"Java Script"として解釈実行する。

【0023】なお、"Netscape Navigator"以外でも、同等の機能を持つブラウザにも本発明は適用可能である。また、上記の実施例では、ハイパーテキストとしてHT

MLを、コマンドとして"Java Script"を対象として説明したが、この例に限定されることなく任意のハイパーテキスト及びコマンドで同様の処理が可能である。

【0024】なお、本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々変更・応用が可能である。

【0025】

【発明の効果】上述のように、本発明のハイパーテキスト音声制御方法及び装置によれば、ハイパーテキスト中の記述に応じて自動的に制御動作を変える音声制御機能を、ハイパーテキスト表示装置に付加することができる。また、先に出願した特願平8-312016『ハイパーテキスト中継方法及び装置』では、ハイパーテキストを中継する際にハイパーテキストの内容を自動的に解析し、その内容に応じて音声認識タグ及び発音表記とコマンドの組からなるリストを当のハイパーテキストに挿入する処理を行うが、この方法を本発明と組み合わせることにより、任意のハイパーテキストに対して、リンクと結びついたテキストを利用者が発話することで、そのリンク先にハイパーテキスト表示装置の表示を切り替えることが可能となる。

【図面の簡単な説明】

【図1】本発明の原理を説明するための図である。

【図2】本発明の原理構成図である。

【図3】本発明のハイパーテキスト音声制御装置の構成図である。

【図4】本発明のハイパーテキスト音声制御動作を示す図である。

【図5】本発明の一実施例のシステム構成図である。

【図6】本発明の一実施例の音声認識を可能とするHTMLの記述例である。

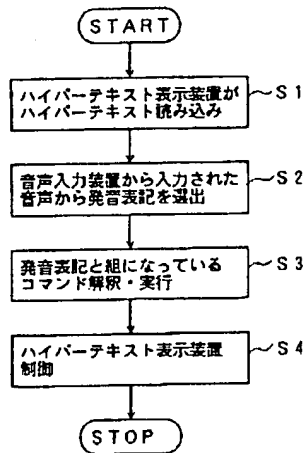
【符号の説明】

- 10 コンピュータネットワーク
- 20 ハイパーテキスト表示装置
- 30 音声入力装置、音声入力手段
- 40 音声認識装置、音声認識手段
- 50 コマンド解釈実行装置、コマンド解釈実行手段
- 100 ブラウザ
- 101 HTML表示装置
- 102 Java Script 解釈実行装置
- 110 音声認識装置
- 120 音声入力装置
- 130 マイク

BEST AVAILABLE COPY

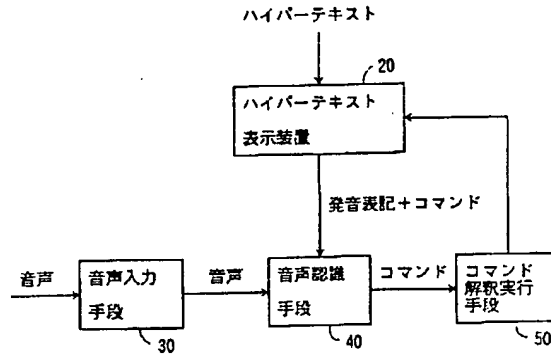
【 図1 】

本発明の原理を説明するための図



【 図2 】

本発明の原理構成図

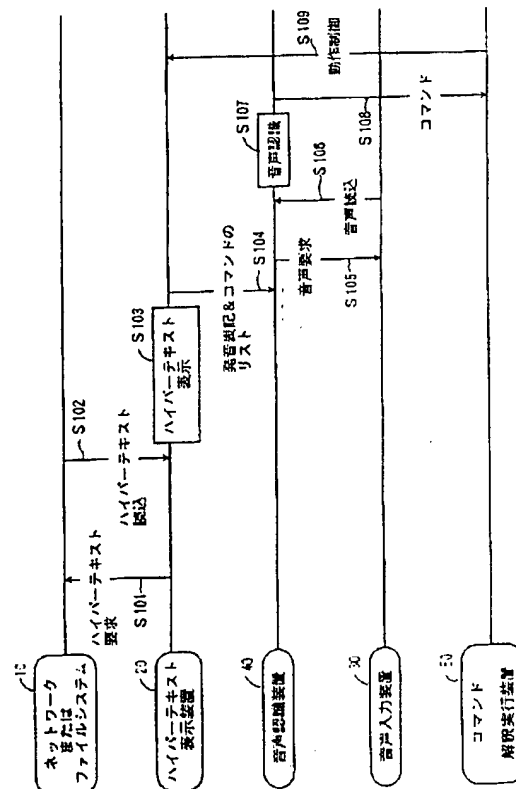
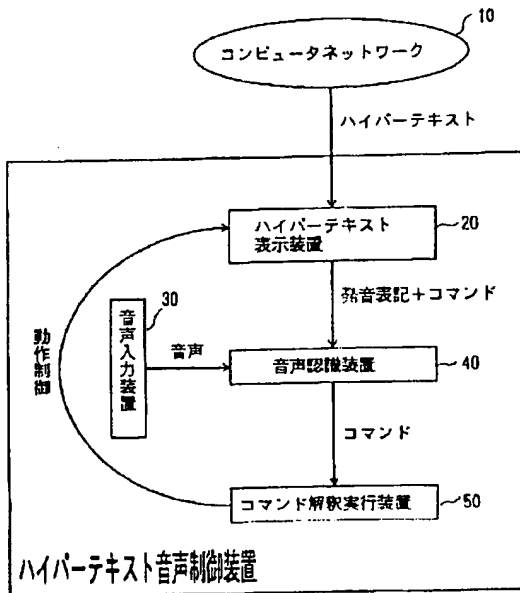


【 図4 】

本発明のハイパーテキスト音声制御動作を示す図

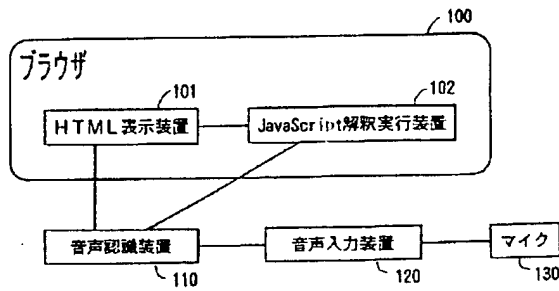
【 図3 】

本発明のハイパーテキスト音声制御装置の構成図



【 図5 】

本発明の一実施例のシステム構成図



【 図6 】

本発明の一実施例の音声認識を可能とするHTMLの記述例

```

<HTML>
<BODY>
リンクのページ<BR>
<A HREF="http://www.com/fruit.html">果物</A><BR>
<A HREF="http://www.com/animal.html">動物</A><BR>
<A HREF="http://www.com/plant.html">植物</A><BR>
<A HREF="http://www.com/building.html">建物</A><BR>
<APPLET CODEBASE="http://voica.ntt.co.jp" CODE="VoiceRecognition.class">
<PARAM NAME="1" VALUE="くだもの.window.open('http://www.com/fruit.html')">
<PARAM NAME="2" VALUE="どうぶつ.window.open('http://www.com/animal.html')">
<PARAM NAME="3" VALUE="しょくぶつ.window.open('http://www.com/plant.html')">
<PARAM NAME="4" VALUE="たてもの.window.open('http://www.com/building.html')">
</APPLET>
</BODY>
</HTML>

```

BEST AVAILABLE COPY